DOI:
10.1021/acs.biochem.8b01101

Document Version
Peer reviewed version

Citation for published version (APA):
Larocca, M., Foglia, F., & Cilibrizzi, A. (2019). Dihedral Angle Calculations to Elucidate the Folding of Peptides through Its Main Mechanical Forces. *Biochemistry*, *58*(8), 1032-1037.
https://doi.org/10.1021/acs.biochem.8b01101

Dihedral angle calculations to elucidate the folding of peptides through its main mechanical forces

Journal:	*Biochemistry*
Manuscript ID	bi-2018-011014.R2
Manuscript Type:	From the Bench
Date Submitted by the Author:	n/a
Complete List of Authors:	Larocca, Michele; Università di Bologna, Dipartimento di Chimica Ciamician Foglia, Fabrizia; King's College London, Institute of Pharmaceutical Science Cilibrizzi, Agostino; King's College London, Institute of Pharmaceutical Sciences

Dihedral angle calculations to elucidate the folding of peptides through its main mechanical forces

Michele Larocca, Fabrizia Foglia and Agostino Cilibrizzi[a]*

[a]Institute of Pharmaceutical Science, King's College London, Stamford Street, London SE1 9NH, UK.

Corresponding Author

Dr Agostino Cilibrizzi

Institute of Pharmaceutical Science, King's College London

5th Floor, Franklin-Wilkins Building, 150 Stamford Street

London SE1 9NH

Tel +44 (0)207 848 4501

E-mail agostino.cilibrizzi@kcl.ac.uk

ABSTRACT

This study reports a general method to calculate dihedral angles (φ and ψ) of a given amino acid sequence, focusing on potential energy and torque moment concepts. By defining these physical measures in relation to the chemical interactions that occur on each single amino acid residue within a peptide, the folding process is analysed as the result of main mechanical forces (MMF) exerted in the specific amino acid chain of interest. As a proof of concept, Leu-enkephalin was initially used as a model peptide to carry out the theoretical study. Our data show agreement between calculated Leu-enkephalin backbone dihedral angles and the corresponding experimentally determined x-ray values. Hence, we used calcitonin to validate our MMF-based method on a larger peptide, i.e. 32 amino acid residues forming an α–helix. Through a similar approach (although simplified with regards to electrostatic interactions), the calculations for calcitonin also demonstrate a good agreement with experimental values. This study offers new opportunities to analyse a peptides' amino acid sequences and help in the prediction of how they must fold, assisting the development of new computational techniques in the field.

KEYWORDS: polypeptide folding; dihedral angles; amino acid sequence; mechanical forces; potential energy; torque moment; Leu-enkephalin; calcitonin.

Introduction

Protein folding has been widely investigated over the years providing significant advances in the study of specific amino acid sequence roles and their intrinsic driving forces that are crucial for the process.[1-12] To date, the thermodynamic hypothesis provides a fundamental output with regards to the information required to fold a polypeptide chain.[13] In addition, the role played by the dominant forces involved in the folding of a protein, such as hydrophobic effect, hydrogen bonding and configurational entropy,[1-3,14-16] is also an essential aspect.

The interactions among residues in a chain showed to occur through the formation of hydrophobic cores, hydrogen bonds, electrostatic interactions, π-π interactions, C-H··π-cloud interactions and steric repulsion (from forbidden overlaps), which contribute to the folding of a specific polypeptide in its most stable three-dimensional structure.[17-19] This process is strictly related to the energy landscape explored by the amino acid sequence during folding, so that the sequence itself can fold correctly and rapidly. Nevertheless, how the amino acid sequence encodes a unique and specific energy landscape is still an unanswered question in the field.[7-12,20-22]

To date, computational biophysical studies on macromolecular systems are based on molecular dynamic simulations set in the conditions of interest, having geometries and interacting parameters developed separately from the macromolecular system under study. In this way, the parameters used to describe interactions do not show accurate,[12] although the transfer across different systems might represent an advantage. Limits arise from the presence of several of atoms which, in turn, produce many interactions, rendering the prediction of protein structure likely impossible.[5] Clearly, new approaches are required in order to balance the computational effectiveness and the real physical context of a specific macromolecular system.[12]

In this context, by considering the possible interactions among each single amino acid residue of a polypeptide chain, the present study aims to define the exerted main mechanical forces (MMF) and calculate the backbone dihedral angles by mean of empirical potential energy functions, in

order to ultimately produce a general physical approach for a more detailed elucidation of the whole folding process. Indeed, the number of rotamers for the specific backbone, also represents a crucial determinant for the folding process, due to the defined structural restrictions arising from local interactions of conformational isomers. Specifically, the latter affects the stable interplay in the amino acid chain that can act as nucleation point leading to the correct folding of a particular sequence.[23,24]

In this view, we envisaged rationalising the main interactions and mechanical forces occurring among the single residues of an amino acid sequence by considering the concepts of potential energy and torque moment (or twisting), which can be defined by a systematic analysis of the amino acid sequence. For our purpose, we firstly selected Leu-enkephalin (a small endogenous oligopeptide with agonistic activity on μ- and δ-opioid receptors) as a prototype to carry out the theoretical study. Each function, charged group and substituent in Leu-enkephalin has been "handled" as element able to produce a specific interaction that can be described and quantified by an established mechanical force. In turn, this provides the necessary rotation values to obtain the related backbone dihedral angles. In the method, two interacting elements need to be considered as a two-member system (e.g. two-rings = two π-clouds) or two-charge system (i.e. dipole), in order to define the equations that relate to potential energy and torque moment of the system. In turn, these empirical potential energy functions allow to calculate the dihedral angle or the partial dihedral angle. Being the hydrophobic effect considered as the main driving force for the folding of peptides [2,25], we have identified the π–π interaction between the aromatic rings of L-Tyr1/L-Phe4 and the C-H···π-cloud interaction between the L-Phe4/L-Leu5 side chains as the main interactions between apolar residues (i.e. as the expression of the hydrophobic effect).

Furthermore, we have studied a larger peptide – i.e. calcitonin, in order to widely validate the proposed physical-chemistry approach. As reference, we have adopted the NMR-solved structure of the so-reported 'conformer one' for calcitonin,[26] which is referred as the most representative term of the ensemble. Due to the increased size of this second peptide, the calculations for the

electrostatic interactions were performed through a slightly simplified method with regards to the one used for Leu-enkephalin. Specifically, in this case we have considered the sin of the Coulomb force and the force as the product of an attraction between the charges, in order to determine the lever arm of the torque moment (whereas the sin component of the distance vector, with the Coulomb force exerted singularly by the charge on the α–carbons, was considered for the torque moment in Leu-enkephalin). Clearly, our results demonstrate that both methods can be efficiently employed to calculate the torque moment. Moreover, for the hydrophobic effect in calcitonin, we have adopted the same approach of Leu-enkephalin and an additional electrostatic interaction between apolar residues (i.e. the so-called 'hydrophobic effect' or 'hydrophobicity'), which is considered as the output of London dispersion forces.

With this approach, the calculated Leu-enkephalin and calcitonin backbone dihedral angles showed a robust agreement with the values experimentally obtained by x-ray diffraction[27] and NMR,[26] respectively.

Experimental Details

Structural model: the theoretical basis used to calculate the backbone dihedral angles φ and ψ of Leu–enkephalin.

The primary structure of Leu–enkephalin was analysed in a fully extended conformation (i.e. primary structure, Figure 1) by adopting previously reported bond lengths and molecular geometry as reference geometrical parameters.[25] The following interactions and mechanical forces were considered to drive the folding process: i) Coulomb interaction between the two terminal charges; ii) π–π interaction between two aromatic rings; iii) C-H···π-cloud interaction and iv) proton-proton repulsion due to a forbidden overlap at level of the NH protons of the amide bonds between Gly2 and Gly3. The structural model of Leu-enkephalin is shown in Figure 1 and the measured distances for the calculations are reported in Table 1. Full details of calculations and equations are reported in SI (Sections 1-11).

Table 1. Distances used for the calculations converted in meters and measured in Angstrom from the primary structure of Leu-enkephalin. r_c : distance between the terminal charges, r_+ : distance between the positive charge and the α-carbon of the residue, r_- : distance between the negative charge and the α-carbon of the residue, r_r : distance between the two aromatic rings, r_h : distance between the farther C-H bond of the Leu5 side chain and the aromatic ring of Phe4, r_{LJ}: distance between NH^1 (Gly2) and NH^2 (Gly3). To note here, that each distance is considered as averaged value between the two possible extremes.

Residue Name	Dihedral Angle	$r_c \cdot 10^{-10}$ /m	$r_+ \cdot 10^{-10}$ /m	$r_- \cdot 10^{-10}$ /m	$r_r \cdot 10^{-10}$ /m	$r_h \cdot 10^{-10}$ /m	$r_{LJ} \cdot 10^{-10}$ /m
Tyr1	ψ^1	17.3 ± 0.7	1.46	16.2 ± 0.6	12.7 ± 1.3	-	-
Gly2	φ^2 ψ^2	17.3 ± 0.7	4.86	12.4 ± 0.7	-	-	-
Gly3	φ^3 ψ^3	17.3 ± 0.7	8.52	9.1 ± 0.5	-	-	4.4
Phe4	φ^4 ψ^4	17.3 ± 0.7	11.98	5.4 ± 0.5	12.7 ± 1.3	9.5 ± 0.9	-
Leu5	φ^5	17.3 ± 0.7	15.7	2.4 ± 0.05	-	9.5 ± 0.9	-

Structural model: the theoretical basis used to calculate the backbone dihedral angles φ and ψ of calcitonin.

The primary structure of calcitonin was analysed in a fully extended conformation (i.e. primary structure) from a dedicated pdb file. The following interactions were considered to drive the folding process: i) Coulomb interaction between the two terminal charges; ii) π–π interaction between two aromatic rings; iii) C-H···π-cloud interaction, iv) Van der Waals proton-proton repulsion, v) H-bond and vi) Hydrophobic effect or hydrophobicity – *"London Dispersion Forces"*. Full details of calculations are reported in SI (Section 12).

Results and discussion

This study presents an innovative physical approach to calculate the backbone dihedral angles of a given polypeptide where all the main possible interactions among single residues, as well as derived mechanical forces, are considered as determinants that drive the folding of the chain. As a proof of concept, we have chosen Leu-enkephalin and used as reference its x-ray determined backbone dihedral angles.[27] For our purpose, we have initially created a pdb file of Leu-enkephalin (see section 2, SI) to determine the distances between the various interacting elements in the fully extended molecule (Figures 1 and S1), corresponding to its primary structure – i.e. in its unfolded state and without secondary structure assigned. In this regard, peptides can adopt various conformations, often trapped at a local minimum of the energy landscape, which in turn could show differences in terms of dihedral angle values. With the aim to reduce the variability arising from this, we have used the only allowed fully extended conformation to gain a more accurate calculation/prediction of the dihedral angles.

Our analysis starts on the assumption that what drives the folding process (and, therefore, needs to be taken into account) are the related main mechanical forces in combination with the following interactions: i) electrostatic interaction between the terminal charges, ii) π–π interaction between the aromatic rings of L-Tyr1 and L-Phe4, iii) H-H repulsive interaction (NH-NH forbidden overlap) and iv) C-H···π-cloud interaction between the L-Phe4 and L-Leu5 side chains (Figure 1 and Table 2).

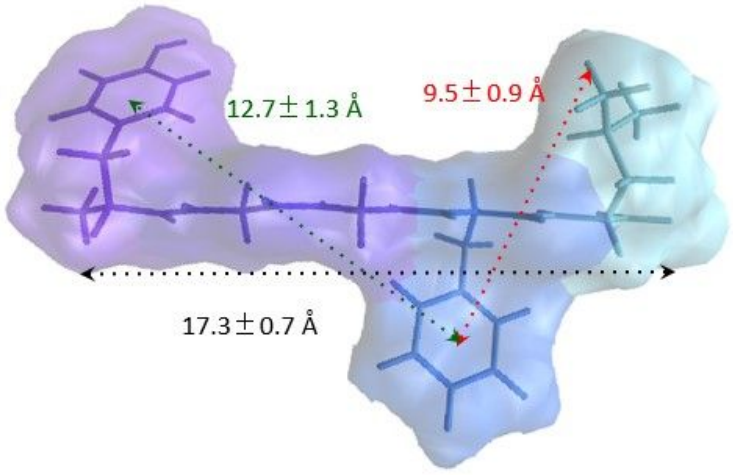

Figure 1. Cartoon representing the primary structure of Leu–enkephalin; each amino acid is represented in different colour from purple (Tyr1) to cyan (Leu5). In the figure are reported only three distances considered for the calculations: i) length of the fully extended peptide ($17.3_{0.7}$ Å; reported in black); ii) π–π interaction ($12.7_{1.3}$ Å; reported in green); iii) *C-H···π-cloud* interaction ($9.5_{0.9}$ Å; reported in red).

Table 2. Comparative list of backbone dihedral angles in Leu-enkephalin, showing the agreement between experimentally determined [27] and theoretically calculated (including their standard deviation) values.

Residue Name	Dihedral Angle	Experimental Values[27]	Calculated Values
L-Tyr1	ψ^1	126°	129.62 ± 0.15°
L-Gly2	φ^2	59°	58.77 ± 0.37°
L-Gly2	ψ^2	25°	25.89 ± 0.22°
L-Gly3	φ^3	97°	97.22 ± 0.25°
L-Gly3	ψ^3	-7°	-6.69 ± 0.61°
L-Phe4	φ^4	-136°	-137.21 ± 1.12°
L-Phe4	ψ^4	145°	150.79 ± 2.21°
L-Leu5	φ^5	-105°	-106.35 ± 1.23°

For each set of calculations, the effect of the considered force was distinguished in two parts: i) determination of the force itself (F) together with the potential energy of the system (U); ii) effect of the force on the α-carbon of the dihedral angle of interest. The latter only accounts for the component of the force able to induce a rotation (Fsin(x)). This approach allows the evaluation of the torque moment of the force (τ) acting on the α-carbon, which is believed to drive the rotation in order to obtain the desired dihedral angle. By comparing U and τ, it was therefore possible to calculate a total or partial backbone dihedral angle. In this regard, it was also needed to define the distances in the primary structure by introducing a folding (or twisting) factor (θ), being fundamental to obtain the dihedral angle values. The folding factor accounts for the repositioning of a specific element (side chain, charge or NH bond) while folding, which is directly connected to the molecular interaction involving the residue of interest. In turn, the folding factor allows to quantify the related changes in distance. It is also worth to consider the role of the terminal charges which have shown to affect the calculation of all the dihedral angles. In this regard, in the present study we report two alternative methods, since we have also developed a simplified calculation approach that takes into account electrostatic interactions (*vide infra*, for calcitonin). Specifically, in each residue, one of the partial dihedral angle arises from the electrostatic interaction that can induce the following partial twisting of the considered bond, by acting on the α-carbon. This factor was then added to the further partial twisting(s) to allow accurate calculation of the final dihedral angle(s).

Starting from Tyr1, ψ^1 (Eq. A; section 4, SI) depends on two partial twisting contributions arising from: i) the two terminal charges (ψ^e; section 4.1, SI) and ii) from the π–π interaction (ψ^p; section 4.2, SI). Overall, the charges were considered as a dipole with relative potential energy (U_e; Eq. S4, SI) to be determined. A similar approach has been also adopted for all ψ^e and φ^e calculations in the whole study. Noteworthy, the distance r_e (needed here to calculate the Coulomb force and therefore the dipole potential energy) is not affected by any folding factor. Therefore, the peptide can be considered as "fully denatured". Furthermore, the torque moment (τ_e; Eq. S8, SI) was also

calculated, in order to estimate the effect of single charges on a point in between the dipole itself (i.e. always the α-carbon). In turn, the torque moment requires the knowledge of the relative distance between the charge and considered α-carbon. This translates into a value of $44.43_{0.15}°$ (for details of calculations see Table S1 and section 4, SI).

As mentioned above, ψ^1 also depends upon π-π interaction between Tyr1 and Phe4 side chains (ψ^p). In this case the calculation of the force accounts for the distance between the two aromatic rings (r_r, Figure S3) together with the folding factor θ^p (Eq. S15, SI). The latter was used to estimate the change in distance between the rings as arising from the rotation around their sp^3 β-carbons. This molecular feature was considered when calculating the dihedral angles ψ^1, φ^4 and ψ^4 that are strictly related to the π–π interaction (see sections 4-6 in SI). In all the three cases, the folding factors are related to an angle of 109.50° (i.e. tetrahedral molecular geometry), which is expressed accounting for the degrees of freedom (i.e. 1, in this particular case). Furthermore, the equation for the torque moment (τ_p; Eq. S18, SI) needs also to account for a final equilibrium distance ($r_f = 4.00 \cdot 10^{-10}$ m; which is the requirement for a π-π interaction to take place).[28] This translates into a value of 85.19° (for details of calculations see Table S1 and section 4, SI). Therefore, by summing ψ^e and ψ^p, the total dihedral angle ψ^1 is $129.62_{0.15}°$ (section 4, SI), which is in close agreement with the experimentally determined value (i.e. 126°).[27]

As expected, the role of π-π interaction needs also to be considered for Phe4 dihedral angle φ^4 (section 5, SI). In line with our approach, this dihedral angle must account for four terms: i) partial twisting from the two terminal charges (φ^e; section 3.1, SI), ii) partial twisting from the π–π interaction (φ^p; section 3.2, SI), iii) partial twisting due to a C-H···π-cloud interaction (φ^h; section 3.3, SI) and iv) a factor to convert the angle into the corresponding counterclockwise value (-360°; Eq. B, section 5, SI). With regards to φ^e, a folding factor (θ^e) is needed, in order to estimate the real distance between the charges of the dipole:

$$\theta^e = \eta_{\psi 1}\,[\pi \cdot r_c \cdot \sin(129.51°)] + \eta_m\,[\pi \cdot r_c \cdot \sin(109.50°)] + \eta_{mh}\,[\pi \cdot r_c \cdot \sin(109.50°)] = 4.09 \cdot 10^{-8}\,\text{m} \quad (1)$$

In this equation, the folding factor (θ^e) depends upon the distance between the charges of the dipole, which in turn showed to relay on the distance (r_c) from the primary structure, together with: i) the tendency to have π-π interaction and ii) the free motion around a sp^3 carbon. The latter is considered in terms of: i) co-planarity which translates into an influence either on the position between the charges of the dipole (e.g. Gly3-Phe4) or on to free rotation of the whole plane itself (i.e. Gly3-Gly2) and ii) tendency to establish a C-H···π-cloud interaction (e.g. by considering only Leu5 side chain). To further express the torque moment, the position of the single charges is considered together with the force acting on Phe4 α-carbon. Although this is an analogous approach as the one defined by Eq.1, the effect of the single charge needs here an extra tailoring. For instance, the folding factor θ^+ depends on the first two terms, while the folding factor θ^- depends only on the third term. The result is a partial twisting of $50.33_{1.12}°$ from the dipole (Table S2 and Eq. B, section 5, SI). Moreover, we have also calculated the partial twisting from both the π-π interaction and the C-H···π-cloud (section 5.3, SI). The resulting final value for φ^4 is $222.79_{1.12}°$. This result exceeds the maximum allowed rotation, according to the partial double bond character of the peptide bond (i.e. maximum allowed clockwise rotation of ±180°). Therefore, by converting the obtained angle into the corresponding counterclockwise value (i.e. addition of -360°), $-137.21_{1.12}°$ is obtained for φ^4 (Eq. B, section 5, SI), in agreement with -136° (i.e. literature value).[27]

Similarly, ψ^4 depends upon three partial twisting: i) from the π–π interaction (involving Phe4 side chain, as in the case of ψ^1 and φ^4; section 3.2, SI); ii) from the C-H···π-cloud interaction involving Phe4 and Leu5 (as in the case of φ^4; section 3.3, SI) and iii) from the two terminal charges (Eq. C, section 6, SI). The latter accounts for a folding factor (θ^e), which affects the distance between the charges:

$$\theta^e = \eta_{\varphi 4} [\pi \cdot r_c \cdot \sin(-136.41°)] = 7.18 \cdot 10^{-9}\ m \quad (2)$$

As a result, the position of each single charge was calculated with respect to the α-carbon. This allows the determination of the torque moment (τ_e), considering that the positive charge is not

affected by any folding factor. In contrast, the position of the negative charge showed to depend on the π–π interaction. The calculation translates into a value of $150.79_{2.21}°$ for ψ^4.

The α-carbon of Gly2 is also co-planar with the one belonging to Tyr1. This feature is crucial to calculate the dihedral angle of interest φ^2 ($58.77_{0.37}°$; Eq. D, section 7, SI), which is defined only by the electrostatic interaction between the terminal charges (i.e. no side chains are present to produce further interactions). Specifically, this co-planarity also affects the position of the single charges with respect to the α-carbon itself, thus the influence of the π-π interaction plays a role to fix the distance between the positive charge and the α-carbon. The co-planarity between Gly2 and Gly-3 α-carbons is crucial to determine ψ^2 ($25.89_{0.22}°$; Eq. E, section 8, SI), by defining a plane that might have a free rotational motion itself and, therefore, it could explain the absence of any folding factor for the dipole and the positive charge. In this scenario, the torque moment shows only affected by the folding factor θ^-. The same applies to ψ^3 (Eq. F, section 9, SI).

In the case of φ^2, ψ^2 and ψ^3, solely the effect of the electric dipoles on their α-carbon (namely the twisting arising from terminal charges) has been considered, hence the partial dihedral angles correspond to the dihedral angles themselves. It is additionally worth to note that these three angles belong to the two Gly residues (Gly2 and Gly3), which do not possess side chains involved in further molecular interactions. Accordingly, it can be argued that particular care needs to be given to the distance between elements involved in molecular interactions, although terminal charges are fundamental determinants. Indeed, distances depend on the twisting, which arises from the interactions themselves, and on the co-planarity of the atoms lying on the amide bond plane, which plays an important role mainly in relation to α-carbons. This indicates that the dihedral angle related to the N-Cα can be affected by the π–π interaction together with the partial twisting from the terminal charges, with regards to a plane involving the Cα-C bond of the residue 'i' (where the dihedral angle arises from a partial twisting due to a π–π interaction) and the N-Cα bond of the residue 'i+1'. Furthermore, the π–π interaction needs to be additionally considered either as a partial dihedral angle (e.g. ψ^1, φ^4 and ψ^4) or as part of the folding factors used to determine the

intermolecular distances required for the new angle to be calculated. In the latter case, the π–π interaction was considered to define the folding factors related to the calculation of the distance between the terminal charges (dipole) and to calculate the distance between each single charge and the α-carbon of interest.

Moving forward, in the case of φ^3 (Eq. G, section 10, SI), we need to take into account the electrostatic interaction from the terminal charges (which will be treated as above) together with a proton-proton repulsion (φ^{LJ}: partial twisting from amide protons forbidden overlaps – Lennard-Jones potential; section 3.4, SI). The latter accounts for an intra-chain repulsion from amide protons forbidden overlap. This molecular feature relies on considering the progress about the folding of the backbone which leads to a local constraint, hence a further rotation (27.28°; Table S7, SI) arising from the repulsive force exerted. Specifically, the NH group involved in the NH-Cα bond has been regarded as the only functional group able to provide an interaction (since there is no side chain for Gly3) and further produce a torsion. Considering this NH is not involved in hydrogen bonding (i.e. due to the number of residues in the peptide), the forbidden overlap is the only possible interaction occurring. By adding together the two contributions, we obtain a value of $97.22_{0.25}$° for φ^3 (Eq. G and Table S7, section 10, SI). A similar approach (as above) has been adopted to determine φ^5 (Eq. H, section 11, SI). More specifically, in this latter the electrostatic term and the partial twisting from a C-H···π-cloud interaction have been considered, while only the electrostatic term was taken into account in the previous case.

To fully validate the physical-chemistry approach, we tested our MMF-based method on calcitonin, as representative example of a polypeptide. In this case, additional interactions also need to be considered in the calculations, such as H-bonds and hydrophobicity (section 12.1.6, SI), as well as an alternative method to handle the electrostatic interactions (section 12.1.1, SI). The calculated dihedral angles were compared to those of conformer one, previously established as the most representative isomer of calcitonin NMR ensemble.[26] The agreement is high between NMR-determined and calculated dihedral angles (section 12.2, SI), with only four residues that possess a

variable degree of difference between calculated values and the corresponding experimental data (i.e. ψ^4, ψ^{10}, ψ^{28} and ψ^{30}). This is reasonable due to the higher complexity of the structure, with regards to dihedral angle determinations. However, the values obtained through the present theoretical approach are typical α-helix values, in agreement with the reported secondary structure of calcitonin. Moreover, the structure of calcitonin that we have used as reference is a single NMR conformer, out of an ensemble of hundred possible conformers. [26] This accounts for a large variation of possible results, which in our method reflects for only four dihedral angles not matching the experimental data for the selected isomer. Figure 2 shows the superimposition for calculated and NMR-determined structures, demonstrating that only minimal differences are present between the two outputs.

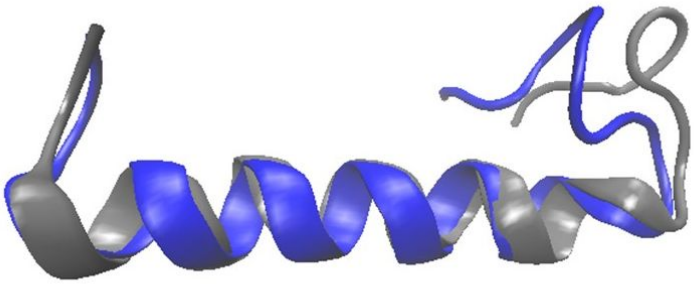

Figure 2. Superimposition for calculated (grey) and NMR-determined (blue) structures of calcitonin.

In this study, Leu-enkephalin has been initially used as a model peptide to evaluate the role of amino acid interactions in the folding of peptide chains. Secondly, calculation tests have been also performed on calcitonin to widely validate the MMF-based method.

By focusing on the chemical interactions among single amino acid residues and the exerted

mechanical forces, a theoretical method has been produced to calculate the backbone dihedral angles. These results express a close matching with experimentally determined values, [26,27] confirming the hypothesis that interactions among the residues might drive the folding process of such a polypeptide chain, by acting through their related mechanical forces which are defined by mean of empirical potential energy functions.

Noteworthy, Anfinsen's dogma suggests that the active structure of a protein, at ambient conditions, is primarily determined by the sequence of its residues. However, this seems valid only for small proteins.[32] In contrast, in the case of polypeptides (i.e. >100 residues) more than one intermediate structure is found during the entire folding process, suggesting that more than one possible "folding pathway" could take place.[33-35] Nevertheless, different 'folding paths' could indeed represent a limit for our calculations, especially in the case of large globular proteins. In this regard, further developments are needed to establish a theoretical approach which would unequivocally determine the values of the degrees of freedom for the twisting factors, in order to systematically define the specific folding factors for each postulated interaction.

Overall, the method described herein shows new possibilities to predict dihedral angles by combining supposed chemical interactions and related empirical potential energy functions, such as the torque moment. Our calculations and data represent a good starting point to evaluate the role of single intra-chain amino acid interactions, which can be fully examined and clarified by considering the spontaneous folding of a particular polypeptide chain. The method allows the prediction of dihedral angles, to help in deciphering the whole folding process for proteins whose crystal structure is unavailable. Specifically, the calculations reported herein represent a useful tool to assist both experimental (e.g. NMR spectroscopy and x-ray crystallography) and computational (e.g. to establish efficient algorithms for new homology models) techniques in the field,[29-31] in order to ultimately calculate and/or predict how a protein amino acid sequence must fold.

ASSOCIATED CONTENT

Supporting Information. Detailed approach of calculation used in this study to determine dihedral angles of Leu-enkephalin and calcitonin.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ACKNOWLEDGMENT

A.C. greatly acknowledges Theragnostics Ltd (https://theragnostics.com/) for financial support to his research.

Declarations of interest: none.

REFERENCES

[1] Camilloni, C., Bonetti, D., Morrone, A., Giri, R., Dobson, C. M., Brunori, M., Gianni, S. and Vendruscolo, M. (2016) Towards a structural biology of the hydrophobic effect in protein folding. Sci. Rep. 6, 28285.

[2] Baldwin, R. L. (2014) Dynamic hydration shell restores Kauzmann's 1959 explanation of how the hydrophobic factor drives protein folding. *Proc. Natl. Acad. Sci. USA 111*, 13052-13056.

[3] Perunov, N. and England, J. L. (2014) Quantitative theory of hydrophobic effect as a driving force of protein structure. *Prot. Sci. 23*, 387-399.

[4] Li, H., Tang, C. and Wingreen, N. S. (1997) Nature of Driving Force for Protein Folding: A Result From Analyzing the Statistical Potential. *Phys. Rev. Lett. 79*, 765.

[5] BakerD. (2000) A surprising simplicity to protein folding. *Nature 405*, 39-42.

[6] Dill, K. A., Bromberg, S., Yue, K., Fiebib, K. M., Yee, D. P., Thomas, P. D. and Chan, H. S. (1995) Principles of protein folding — A perspective from simple exact models. *Prot. Sci. 4*, 561-602.

[7] Pace, C. N., Scholtz, J. M., and Grimsley, G. R. (2014) Forces stabilizing proteins. *FEBS Lett. 588*, 2177-2184.

[8] Baldwin, R. L. (1994) Matching speed and stability. *Nature 369*, 183-184.

[9] Chekmarev, S. F. (2015) Protein Folding as a Complex Reaction: A Two-Component Potential for the Driving Force of Folding and Its Variation with Folding Scenario. *PLoS ONE 10*, e0121640.

[10] Chekmarev, S. F. (2013) Protein folding: Complex potential for the driving force in a two-dimensional space of collective variables. *J. Chem. Phys. 139*, 145103.

[11] Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol. 4*, 10-19.

[12] Ding, F., and Dokholyan, N. V. (2005) Simple but predictive protein models. *Trends Biotechnol. 23*, 450-455.

[13] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science 181*, 223-230.

[14] Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry 29*, 7133-7155.

[15] Javadi, Y. Fernandez, J. M., and Perez-Jimenez, R. (2013) Protein folding under mechanical forces: a physiological view., Physiology 28, 9-17.

[16] Baldwin, R. L., and Rose, G. D. (2016) How the hydrophobic factor drives protein folding., Proc. Natl. Acad. Sci. USA 113, 12462-12466.

[17] Szilágyi, A., Kardos, J., Osváth, S., Barna, L., and Závodszky, P. in Handbook of Neurochemistry and Molecular Neurobiology: Neural protein metabolism and function, ed. Lajtha A. and Banik, N. Springer, Boston (MA), 2007, 303-343.

[18] Chothia, C. (1984) Principles that determine the structure of proteins. Ann. Rev. Biochem. 53, 537-572.

[19] Grantcharova, V., Alm, E. J., Baker D. and Horwich, A. L. (2001) Mechanisms of protein folding. *Curr. Opin. Struct. Biol. 11*, 70-82.

[20] Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884-890.

[21] Dill, K. A. (1999) Polymer principles and protein folding. *Protein Sci*. *8*, 1166-1180.

[22] "Go with the fold" in Chemistry World, pp. 24-27 (printed ed) or under https://www.chemistryworld.com/feature/go-with-the-fold/3008748.article (accessed April 2018).

[23] Levinthal, C. in Mössbauer Spectroscopy in Biological Systems; Proceedings of a meeting held at Allerton House, Monticello, Illinois, ed. P. Debrunner, J.C.M. Tsibris and E. Münck, University of Illinois Press, Urbana (IL), 1969, 22-24.

[24] Grdadolnik, J., Mohacek-Grosev, V., Baldwin, R. L., and Avbelj, F. (2011) Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proc. Natl. Acad. Sci. USA 108*, 1794-1798.

[25] Kauzmann, W. (1964) The Three Dimensional Structures of Proteins. *Biophys. J. 4*, 43-54.

[26] Andreotti, G., Lopez-Mendez, B., Amodeo, P., Morelli, M.A., Nakamuta, H., and Motta, A. (2006) Structural Determinants of Salmon Calcitonin Bioactivity. *J. Biol. Chem. 281*, 24193-24203.

[27] Ishida, T., Kenmotsu, M., Minot, Y., Inoue, M., Fujiwara, T., Tomitai, K., Kimura T., and Sakakibara, S. (1984) X-ray diffraction studies of enkephalins. Crystal structure of [(4′-bromo) Phe4,Leu5]enkephalin. *Biochem. J. 218*, 677-689.

[28] Hunter, C. A. and Sanders, J. K. M. (1990) The nature of pi.-.pi. interactions. *J. Am. Chem. Soc. 112*, 5525-5534.

[29] Zhou, R. Modeling of Nanotoxicity, Springer International Publishing, Switzerland, 2015, 169-189.

[30] Do, T. D., LaPointe, N. E., Sangwan, S., Teplow, D. B., Feinstein, S. C., Sawaya, M. R., Eisenberg, D. S., and Bower, M. T. (2014) Factors that drive peptide assembly from native to amyloid structures: experimental and theoretical analysis of [leu-5]-enkephalin mutants. *J. Phys. Chem. B 118*, 7247-7256.

[31] Sborgi, L., Verma, A., Piana, S., Lindorff-Larsen, K., Cerminara, M., Santiveri, C. M., Shaw, D. E., de Alba, E., and Muñoz, V. (2015) Interaction networks in protein folding via atomic-

resolution experiments and long-time-scale molecular dynamics simulations. *J. Am. Chem. Soc. 137*, 6506-6516.

[32] Sharma, V., Kaila V., and Annila, A. (2009) Protein folding as an evolutionary process. *Physica A Stat. Mech. Appl. 388*, 851-862.

[33] Roder, H., and Colón, W. (1997) Kinetic role of early intermediates in protein folding**.** *Curr. Opin. Struct. Biol. 7*, 15-28.

[34] Sanchez, I. E., and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J.Mol. Biol*. *325*, 367-376.

[35] Khan, F., Chuang, J. I., Gianni, S., and Fersht, A. R. (2003) The kinetic pathway of folding of barnase. *J.Mol. Biol. 333*, 169-186.

For Table of Contents Use Only